SIAM Review
Vol. 44, No. 1, pp. 95–108


# Perfect Packing Theorems and the Average-Case Behavior of Optimal and Online Bin Packing*

E. G. Coffman, Jr.†
C. Courcoubetis‡
M. R. Garey§
D. S. Johnson¶
P. W. Shor¶
R. R. Weber‖
M. Yannakakis**

**Abstract.** We consider the one-dimensional bin packing problem under the discrete uniform distributions $U\{j,k\}$, $1 \leq j \leq k-1$, in which the bin capacity is $k$ and item sizes are chosen uniformly from the set $\{1,2,\ldots,j\}$. Note that for $0 < u = j/k \leq 1$ this is a discrete version of the previously studied continuous uniform distribution $U(0,u]$, where the bin capacity is 1 and item sizes are chosen uniformly from the interval $(0,u]$. We show that the average-case performance of heuristics can differ substantially between the two types of distributions. In particular, there is an online algorithm that has constant expected wasted space under $U\{j,k\}$ for any $j,k$ with $1 \leq j < k-1$, whereas no online algorithm can have $o(n^{1/2})$ expected waste under $U(0,u]$ for any $0 < u \leq 1$. Our $U\{j,k\}$ result is an application of a general theorem of Courcoubetis and Weber that covers all discrete distributions. Under each such distribution, the optimal expected waste for a random list of $n$ items must be either $\Theta(n)$, $\Theta(n^{1/2})$, or $O(1)$, depending on whether certain "perfect" packings exist. The perfect packing theorem needed for the $U\{j,k\}$ distributions is an intriguing result of independent combinatorial interest, and its proof is a cornerstone of the paper. We also survey other recent results comparing the behavior of heuristics under discrete and continuous uniform distributions.

**Key words.** bin packing, online, average-case analysis, approximation algorithms

**AMS subject classifications.** 68W25, 68W40

**PII.** S0036144501395423

**1. Introduction.** Suppose one is given items of sizes $1,2,3,\ldots,j$, one of each size, and is asked to pack them into bins of capacity $k$ with as little wasted space as possible, i.e., one is asked to find a least cardinality partition (packing) of the set of items such that the sizes of the items in each block (bin) sum to at most $k$. For what

*Published electronically February 1, 2002. This paper originally appeared in *SIAM Journal on Discrete Mathematics*, Volume 13, Number 3, 2000, pages 384–402.
http://www.siam.org/journals/sirev/44-1/39542.html
†Columbia University, New York, NY 10027 (egc@ee.columbia.edu).
‡Department of Computer Science, Athens University of Economics and Business, Athens, Greece (courcou@csi.forth.gr).
§Bell Labs, Murray Hill, NJ 07974.
¶AT&T Labs–Research, Florham Park, NJ 07932 (dsj@research.att.com, shor@research.att.com).
‖Statistical Laboratory, University of Cambridge, Cambridge CB3 0WB, UK (R.R.Weber@statslab.cam.ac.uk).
**Avaya Labs Research, Basking Ridge, NJ 07920 (mihalis@research.avayalabs.com).

values of $j$ and $k$ can the set be packed perfectly (i.e., so that the sizes of the items in each block sum to exactly $k$)? Clearly the sum of the item sizes must be divisible by $k$, but what other conditions must be satisfied? Surprisingly, the divisibility constraint is not only necessary but sufficient. Readers might want to try their hand at proving this. Relatively short proofs exist, as illustrated in the next section, but a certain ingenuity is required to find one. The exercise serves as a warm-up for the following more general and more difficult theorem, also proved in the next section, in which there are $r$ copies of each size for some $r \geq 1$.

THEOREM 1 (perfect packing theorem). *For positive integers $k$, $j$, and $r$, with $k \geq j$, one can perfectly pack a list $L$ consisting of $rj$ items, $r$ each of sizes 1 through $j$, into bins of size $k$ if and only if the sum of the $rj$ item sizes is a multiple of $k$.*

In set-theoretic terms, the question answered by Theorem 1 is an intriguing puzzle in pure combinatorics. But our motivation to work on it came from its relevance to certain fundamental questions about the average-case analysis of algorithms. In particular, consider the standard bin packing problem, in which one is given a bin capacity $b$ and a list of items $L = (a_1, a_2, \ldots, a_n)$, where each $a_i$ has a positive size $s_i \leq b$, and is asked to find a packing of these items into a minimum number of bins.

In most real-world applications of bin packing, as in Theorem 1, the item sizes are drawn from some finite set. However, the usual average-case analysis of bin packing heuristics has assumed that item sizes are chosen according to continuous probability distributions, which by their nature allow an uncountable number of possible item sizes (see [3, 10], for example). The assumption of a continuous distribution has the advantage of sometimes simplifying the analysis and has been justified on the grounds that continuous distributions should serve as reasonable approximations for discrete ones. But there are reasons to ask whether this is actually true. For example, consider the continuous uniform distributions $U(0, u]$, $0 < u \leq 1$, where the bin capacity is 1 and item sizes are chosen uniformly from the interval $(0, u]$, and the discrete uniform distributions $U\{j, k\}$, $1 \leq j \leq k - 1$, where the bin capacity is $k$ and item sizes are chosen uniformly from the set $\{1, 2, \ldots, j\}$. The limit of the distributions $U\{mj, mk\}$, as $m \to \infty$, is equivalent to $U(0, j/k]$ (after scaling by dividing the item sizes and bin capacities by $mk$). However, in the limit, combinatorial questions such as those addressed by Theorem 1 evaporate. This suggests that something important (and interesting) may in fact be lost in the transition from discrete to continuous models. The results in this paper show that this is indeed the case.

To describe the results, we need the following notation. If $A$ is a bin packing algorithm and $L$ is a list of items, then $A(L)$ is the number of bins used when $A$ is applied to $L$, and $s(L)$ is the sum of the item sizes in $L$ divided by the bin capacity. Note that $s(L)$ is a lower bound on the number of bins needed to pack $L$. The *waste* in the packing of $L$ by $A$ is denoted by $W^A(L) = A(L) - s(L)$. Let $\mathrm{OPT}(L)$ denote an algorithm that always produces an optimal packing. In what follows, $L_n(F)$ will denote a list of $n$ items whose sizes are independent samples from a given distribution $F$. The *expected waste rate* $EW_n^A(F)$ for an algorithm $A$ and distribution $F$ is defined to be the expected value of $W^A(L_n(F))$ as a function of $n$. In what follows we typically abbreviate this as simply the "expected waste." We say a distribution $F$ is a *bounded waste* distribution if $EW_n^{OPT}(F) = O(1)$. As a consequence of Theorem 1 and a classification theorem of Courcoubetis and Weber [11], we can prove the following.

THEOREM 2. *For any $j, k$ with $1 < j < k - 1$, $EW_n^{OPT}(U\{j, k\}) = O(1)$.*

This in itself does not represent a departure from the continuous model, since $U(0, u]$ is also a bounded waste distribution for all $u$, $0 < u < 1$ [3, 16]. The distinction comes when we consider "online" algorithms. In an online algorithm, items are assigned to bins in the order in which they occur in the input list $L$. Each assignment must be made without knowledge of the sizes or number of items later in the list, and once an item is packed it cannot subsequently be moved. This mirrors many practical situations but clearly is a substantial restriction on the power of an algorithm. In particular, we can prove the following.

THEOREM 3. *If $A$ is an online algorithm and $u \in (0, 1]$, then it cannot be the case that $E[W_n^A(U(0, u])] = o(n^{1/2})$.*

In contrast, in the discrete uniform case there is a single online algorithm that has bounded expected waste for all the distributions $U\{j, k\}$, $1 \le j < k - 1$. This is the recently discovered *sum-of-squares* algorithm ($SS$) of [12], defined as follows. Suppose we are packing integer-sized items into bins of capacity $b$. When an online algorithm packs an item $x$ from such a list, the only thing relevant about the current packing is the number $N_h$ of bins whose current contents total $h$, $1 \le h \le b - 1$. $SS$ chooses the bin into which $x$ is to be placed (either a new, previously empty bin or one that is already partially filled but has enough room for $x$) so as to minimize the resulting sum $\sum_{h=1}^{b-1} N_h^2$. Note that $SS$ can be implemented to take $O(b)$ time per item [12], and so runs in linear time for any fixed bin size.

As shown in [12], algorithm $SS$ performs well on average in a surprisingly general sense. Let us say that a *discrete distribution* $F$ is any triple $(b, S, \vec{p})$, where $b$ is an integral bin size, $S = \{s_1, \ldots, s_d\}$ is a finite set of integral item sizes in the range from 1 to $b - 1$, and $\vec{p} = (p_1, \ldots, p_d)$ is a rational *probability vector*, where $p_i > 0$ is the probability of item size $s_i$ and $\sum_{i=1}^{d} p_i = 1$. (We ignore the possibility of items of size $b$ since such items always must start a new bin and completely fill it, leaving the rest of the packing unaffected.) The following specialized version of the result of [12] suffices for our needs.

THEOREM (see [12]). *For any* discrete distribution $F = (b, S, \vec{p})$ *with* $1 \in S$,

$$EW_n^{SS}(F) = \Theta\left(EW_n^{OPT}(F)\right).$$

Hence by Theorem 2, $EW_n^{SS}(U\{j, k\}) = O(1)$ for all $1 < j < k - 1$.

The current paper is organized as follows. The proof of the perfect packing theorem (Theorem 1) appears in section 2. Section 3 then presents the proof of Theorem 2. We begin by describing the classification result of Courcoubetis and Weber [11] upon which the proof depends. This result says that for any discrete distribution $F$, $EW_n^{OPT}(F)$ must be one of $\Theta(n)$, $\Theta(n^{1/2})$, or $O(1)$. Which case applies depends on the existence of certain perfect packings and is in general NP-hard. However, Theorem 1 allows us avoid this complexity in the case of the discrete uniform distributions $U\{j, k\}$. Theorem 3, this paper's contribution to the theory of continuous distributions, is proved in section 4. We conclude in section 5 with a survey of the results that have been proved about the average-case behavior of bin packing algorithms under discrete and continuous distributions. As we shall see, there are other significant differences between the discrete and continuous cases.

## 2. The Perfect Packing Theorem.

We begin our proof of Theorem 1 with three lemmas that list a number of special instances that lead to perfect packing. The first lemma takes care of the special case, $r = 1$.

LEMMA 4. *Suppose $m$, $j$, and $k$ are positive integers such that $j \le k$ and $mk = j(j + 1)/2$. Then the set of $j$ items, one each of sizes* $1, \ldots, j$, *perfectly packs into $m$ bins of size $k$.*

*Proof.* The proof is by induction. Pick $j$ and $k$ and assume the theorem is true for all pairs that are smaller in lexicographic order than $(k, j)$. The theorem is clearly true for $k \le 2$ or $j \le 2$, so assume $k, j > 2$.

If $j > k/2$, then we can start by perfectly packing bins with pairs of items $(j-i, k-j+i)$, $0 \le i < j - k/2$, after which the remaining items are those of sizes $1, \ldots, k-j-1$, plus the item of size $k/2$ if $k$ is even. Since the sum of the sizes of the items that have been packed at this point is a multiple of $k$, the sum of the sizes of the remaining items is also a multiple of $k$. If $k$ is odd, the unpacked items are an instance of $(k, k-j-1)$, with $k-j-1 < j$ and the induction hypothesis applies. If $k$ is even, then $k/2$ divides $j(j+1)/2$ and all remaining items are no larger than $k/2$. Thus the items $1, \ldots, k-j-1$ form an instance of $(k/2, k-j-1)$ and by the induction hypothesis can be perfectly packed into bins of size $k/2$. These half-bins and the item of size $k/2$ can then be combined into bins of size $k$.

Now suppose $j \le k/2$. If $k$ is even, then we have an instance of $(k/2, j)$ and the induction hypothesis applies. If $k$ is odd, first note that $k/2 \ge j$ and $j > 2$ imply $k > j+1$, which together with $mk = j(j+1)/2$ implies $j > 2m$. Thus we can construct $m$ pairs of items each of total size $k' = 2j - 2m + 1$ by combining $j - i$ with $k' - j + i$, $0 \le i \le m-1$. If we place one pair in each of our $m$ bins, we now have $m$ bins with gaps of size $k - k' = k - 2j + 2m - 1$ and items of sizes $1, \ldots, j - 2m$. Because $mk = j(j+1)/2$, the sum of these item sizes must be $m(k-k')$, and so an application of the induction hypothesis to the instance $(k-k', j-2m)$ completes the proof. □

LEMMA 5. *Consider $r > 1$ sets, the $i$th of which consists of $j$ items of consecutive sizes, $\ell_i + 1, \ldots, \ell_i + j$, for some $\ell_i \ge 0$. Suppose either* (a) *$r$ is even or* (b) *$j$ is odd. Then these $rj$ items perfectly pack into $j$ bins of size equal to the sum of the average item sizes in the $r$ groups, i.e., $r(j+1)/2 + \sum_{i=1}^{r} \ell_i$.*

*Proof.* The lemma will follow if we can show that for $\ell_i = 0$, $1 \le i \le r$, and bins of size $r(j+1)/2$ it is possible to pack perfectly the items into the $j$ bins in such a manner that each bin contains exactly one item from each of the $r$ sets.

If $r$ is even, then we simply take two of the sets and pack the $i$th largest item in one set with the $i$th smallest item in the other set, i.e., as the pair $(i, j-i+1)$, $i = 1, \ldots, j$. This fills $j$ bins to level $j+1$. By repeating this $r/2$ times we fill $j$ bins of size $r(j+1)/2$.

If $r$ and $j$ are both odd, then an extra step is required. The idea is first to pack items in triples, one item from each of three sets, such that the sum of each triple is the same. It is easiest to appreciate the construction by considering an example, say $j = 9$. The triples, which each sum to 15, are given in the columns below.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 |
| 8 | 6 | 4 | 2 | 9 | 7 | 5 | 3 | 1 |

In general, the triples are $(i, i + (j+1)/2, j + 1 - 2i)$, $i = 1, \ldots, (j-1)/2$, and $(i, i - (j-1)/2, 2j - 2i + 1)$, $i = (j+1)/2, \ldots, j$. The result of packing these one per bin is to fill all $j$ bins to level $3(j+1)/2$. The number of remaining sets is even and the remaining spaces in the $j$ bins are equal. Thus, the procedure for case (a) can be applied to complete the packing in each bin. □

The following lemma provides part of the induction step used in the proof of Theorem 1.

LEMMA 6. *Consider a quadruple $(k, j, r, m)$ of positive integers such that $k \ge j$ and $mk = rj(j+1)/2$. Then there exists a perfect packing of $r$ copies of $1, \ldots, j$ into $m$ bins of size $k$ if there exists a perfect packing for each lexicographically smaller quadruple of this form, and if any one of the following holds:*

(a) $j \geq k/2$.
(b) *$r$ does not divide $k$.*
(c) *$k$ or $r$ is even.*
(d) $j \leq (r-1)k/2r$.

*Proof.* First, using the arguments of Lemma 4, we demonstrate how to reduce the problem to a smaller instance if (a) holds. If $j \geq k/2$ and $k$ is odd, then we can pack bins with pairs $(j-i, k-j+i)$, $i = 0, \ldots, j-(k+1)/2$. The remaining items, which are of sizes $1, \ldots, k-j-1$, define the smaller instance $(k, k-j-1, r, m')$, where $m' = m - r(2j+1-k)/2$. If $j > k/2$ and $k$ is even, then we can pack bins in the same way, $i = 0, \ldots, j-1-k/2$. The remaining items, which are of sizes $1, \ldots, k-j-1$ and $k/2$, can be packed into bins of size $k/2$ by the induction hypothesis that there exists a perfect packing for $(k/2, k-j-1, r, m')$, where $m' = 2m - r(2j-k)$.

If (b) holds, then $r$ and $m$ must have a common factor $p > 1$ and the problem reduces to the instance $(k, j, r/p, m/p)$.

Now suppose neither (a) nor (b) holds but (c) does. If $k$ is even, then $k/2$ divides $rj(j+1)/2$. Since (a) doesn't hold, $j < k/2$. Thus the problem reduces to a smaller instance in which the bin size is $k/2$. If $r$ is even, then since (b) does not hold, $k$ is divisible by $r$ and so must be even too. Thus the same argument applies.

Finally, for case (d), assume that (a), (b), and (c) do not hold, i.e., $j < k/2$, $r$ divides $k$, and $k$ and $r$ are both odd. Let $r_1 = (r+1)/2$, $k_1 = kr_1/r$ and $r_2 = (r-1)/2$, $k_2 = kr_2/r$. Note that $r_1 + r_2 = r$ and $k_1 + k_2 = k$. The fact that $r$ is odd implies that $r_1$ and $r_2$ are integers. The fact that $r$ divides $k$ implies that $k_1$ and $k_2$ are integers, with $mk_1 = r_1 j(j+1)/2$ and $mk_2 = r_2 j(j+1)/2$. Since by assumption $j < k/2$, we have $k_1 \geq j$, and hence by hypothesis for the instance $(k_1, j, r_1, m)$, we can pack $r_1$ copies of $1, \ldots, j$ into $m$ bins of size $k_1$. Similarly, if also $j \leq k_2$, then we can pack $r_2$ copies of $1, \ldots, j$ into $m$ bins of size $k_2$. Since $k_1 + k_2 = k$ we can combine pairs of bins of sizes $k_1$ and $k_2$ into bins of size $k$. Thus there is a reduction to smaller instances if $j \leq k_2 = (r-1)k/2r$, i.e., if (d) holds. □

*Proof of the perfect packing theorem.* Instances for which the theorem is to be proved are described by the quadruples of Lemma 6. Notice that it would be enough to specify the triple $(k, j, r)$; however, it is helpful to mention $m$ explicitly. The proof of the theorem is by induction on $(k, j, r)$ under lexicographical ordering. By Lemma 4 it is true for $r = 1$. Assume all quadruples that are smaller than $(k, j, r, m)$ can be perfectly packed and $r > 1$. We show there exists a perfect packing of $r$ copies of $1, \ldots, j$ into $m$ bins of size $k$. By Lemma 6, we need only consider the case when $k$ and $r$ are odd, $r$ divides $k$, and $(r-1)k/2r < j < k/2$. Note that in this case $(r-1)k/2r$ is an integer and $k/2r$ is 0.5 more than an integer. We show below that we can perfectly pack all the items of sizes from $j+1-(r-1)k/2r$ through $j$ into bins of size $k$. (Note that the lower bound on this range is greater than 1 because of the above lower bound on $j$.) The theorem then follows because the remaining items form a smaller quadruple, so by the induction hypothesis they can be perfectly packed into bins of size $k$.

To follow the construction below, the reader may find it helpful to consider a specific example. Consider the quadruple $(k, j, r, m) = (165, 77, 5, 91)$. Note that $k$ and $r$ are odd, $r$ divides $k$, and $j$ lies between $(r-1)k/2r = 66$ and $k/2 = 82.5$. We show below how to perfectly pack all items of sizes $12, \ldots, 77$. The remaining items form the smaller quadruple $(165, 11, 5, 2)$.

To pack all items of sizes from $j+1-(r-1)k/2r$ through $j$, we divide the range of item sizes into intervals, i.e., sets of consecutive integers. Each interval is symmetric about a multiple of $k/2r$ and has one of two lengths depending on whether the interval is symmetric about an odd or even multiple of $k/2r$. To form the intervals, we first take the largest interval that is symmetric about $(r-1)k/2r$; this is the interval $[(r-1)k/r-j, j]$. Note that this interval does not include $(r-2)k/2r$ since $j < k/2 = rk/2r$. Next we take the largest interval that can be formed from the remaining items that is symmetric about $(r-2)k/2r$, obtaining the interval $[j-k/r+1, (r-1)k/r-j-1]$. Continuing in this fashion and taking intervals symmetric about further multiples of $k/2r$, we end up with intervals of two kinds. First, there are $(r-1)/2$ intervals centered on even multiples of $k/2r$, with the interval centered on $(r-1-2i)k/2r$ being $[(r-1-i)k/r-j, j-ik/r]$, where $i$ ranges from 0 to $(r-3)/2$. Second, there is an equal number of intervals centered on odd multiples of $k/2r$, with the interval centered on $(r-2i)k/2r$ being $[j-ik/r+1, (r-i)k/r-j-1]$, where $i$ ranges from 1 to $(r-1)/2$. Note that the smallest endpoint is $j-ik/r+1$ for $i=(r-1)/2$, which equals $j+1-(r-1)k/2r$ as claimed above.

For the numerical example above, there are two intervals of each type. Intervals of the first type are $[22, 44]$ and $[55, 77]$; they are of length 23 and symmetric about 33 and 66. Intervals of the second type are $[12, 21]$ and $[45, 54]$; they are of length 10 and symmetric about 16.5 and 49.5. In general, intervals of the first type have odd length $2j-(r-1)k/r+1$ and are symmetric about an even multiple of $k/2r$. Intervals of the second type have even length $k-2j-1$ and are symmetric about an odd multiple of $k/2r$. Our plan is to use Lemma 5 to perfectly pack into bins of size $k$ those items whose sizes lie in intervals of the same type.

We begin by considering all those intervals of the first type. These have odd lengths and they are symmetric about points $ik/r$, $i=1,\ldots,(r-1)/2$. There are $r$ items of each size in each of these intervals. Our strategy is to partition these intervals into groups that satisfy the hypotheses of Lemma 5(b). That is, we arrange for the midpoints of the intervals within each group to sum to $k$. Since the midpoints correspond to the average item sizes for the corresponding intervals, and the number of items in the intervals is odd, Lemma 5(b) implies that we can perfectly pack the items in the intervals of each group. Constructing these groups is a bin packing problem in which the midpoints of the intervals take on the role of item sizes. In what follows we write "items" in quotes when speaking of the midpoints of intervals, possibly normalized, and viewing them as items to be perfectly packed in bins of some required size. In considering intervals of the first type, it is as though we had $r$ "items" of each of the sizes $ik/r$, $i=1,\ldots,(r-1)/2$, and wished to pack them into bins of size $k$. After a normalization that multiplies each item size by $r/k$, this is equivalent to the problem of packing $r$ "items" of each of the sizes $1,\ldots,(r-1)/2$ into bins of size $r$. That is, we have a smaller version $(k', j', r', m')$ of our packing problem, with $j'=(r-1)/2$, $k'=r'=r$, and $m'=r'j'(j'+1)/2k'$. But by the induction hypothesis this means that the desired packing can be achieved. In the example, it is as though we had 5 "items" of sizes 33 and 66 that are to be packed in bins of size 165. Normalizing by a factor of 1/33, this is equivalent to the problem instance $(5, 2, 5, 3)$.

We must now pack items whose sizes lie in intervals of the second type. These intervals are of even length, symmetric about the points $ik/2r$, for $i$ odd and $i = 1,\ldots,r-2$. Again, there are $r$ items of each size in these intervals. As above, we exhibit a reduction to a smaller perfect packing problem. After we multiply item sizes by $2r/k$ the problem is equivalent to perfectly packing $r$ copies of "items" of sizes $1, 3, 5, \ldots, r-2$ into bins of size $2r$. For the example, this is 5 copies of "items" of sizes 1 and 3, to be perfectly packed into 2 bins of size 10. Unfortunately, if the sum of the "item" sizes is an odd multiple of $r$ the "items" cannot be perfectly packed into bins of size $2r$. For this reason, and also because it is convenient to do so even when the sum of the "item" sizes is a multiple of $2r$, we consider perfect packings into bins of sizes $r$ and $2r$. Assume for the moment that $r$ copies of "items" of sizes $1, 3, 5, \ldots, r-2$ can be perfectly packed into bins of sizes $r$ and $2r$. If they are packed entirely into bins of size $2r$, then the number of "items" in each bin must be even (as all "item" sizes are odd), and so Lemma 5 applies and implies that the original items can be perfectly packed into bins of size $k$. On the other hand, suppose a bin of size $r$ is required. The set of "items" that are packed into a bin of size $r$ corresponds to a set of intervals whose midpoints sum to $k/2$. Recall that the intervals are of even length. We divide each such interval into its first half and its second half, obtaining twice as many intervals, whose midpoints now sum to $k$. Now we can again use Lemma 5 to construct the perfect packing.

The final step in the proof is to show that we can indeed perfectly pack $r$ copies of each of the item sizes $1, 3, 5, \ldots, r-2$ into bins of sizes $r$ and $2r$. We shall use a different packing depending upon whether $r = 4\alpha + 1$ or $r = 4\alpha + 3$.

For the case $r = 4\alpha + 1$, the "items" are perfectly packed by the following simple procedure. We begin by packing one bin with $(1, 1, r-2)$, one bin with $(2i+1, 2i+1, r-2i-2, r-2i)$ for each $i = 1, \ldots, \alpha-1$, and one bin with $(2i-1, 2i+1, r-2i, r-2i)$ for each $i = 1, \ldots, \alpha$ (noting that in the final case, when $i = \alpha$, we get three "items" of size $2i + 1 = r - 2i$). This packs four "items" of each size larger than 1, and three "items" of size 1. We can apply this packing $\alpha$ times, leaving us with one "item" of each size larger than 1 and $\alpha + 1$ "items" of size 1. Then we pack one bin with $(1, 2i-1, r-2i)$ for each $i = 1, \ldots, \alpha$. This uses up all the remaining "items" (where $\alpha + 1$ items of size 1 are used because there are two "items" of size 1 when $i = 1$). For the numerical example, in which $\alpha = 1$, this construction says that we should pack 5 copies of 1 and 3 into bins of size 5 and 10 by first packing one bin with $(1, 1, 3)$ and then one bin with $(1, 3, 3, 3)$. This leaves one "item" of size 3 and two of size 1. These perfectly pack into a bin of size 5.

When $r = 4\alpha + 3$ the procedure is very similar to that above. We begin by packing one bin with $(1, 1, r-2)$, one bin with $(2i+1, 2i+1, r-2i-2, r-2i)$ for each $i = 1, \ldots, \alpha$, and one bin with $(2i-1, 2i+1, r-2i, r-2i)$ for each $i = 1, \ldots, \alpha$. As before, this packs four "items" of each size larger than 1, and three "items" of size 1. We apply this packing $\alpha$ times, leaving us with three "items" of each size larger than 1 and $\alpha + 3$ "items" of size 1. Then we pack one bin with $(2i-1, 2i+1, r-2i, r-2i)$ for each $i = 1, \ldots, \alpha$. This leaves us with $\alpha + 2$ "items" of size 1, two "items" of size $2\alpha + 1$, and one "item" of each other size. Finally, as before, we pack one bin with $(1, 2i-1, r-2i)$ for each $i = 1, \ldots, \alpha + 1$, which uses up all remaining items. ∎

**3. Proof of Theorem 2.** Recall the theorem statement: For any distribution $U\{j, k\}$, with $j < k - 1$, $EW_n^{OPT}(U\{j, k\}) = \Theta(1)$.

We rely on a general result of Courcoubetis and Weber [11]. Suppose $F = (b, S, \vec{p})$ is a discrete distribution as defined in section 1. Note that a packing of items with sizes from $S = \{s_1, \ldots, s_d\}$ into a bin of size $b$ can be viewed as a nonnegative integer vector $\vec{c} = (c_1, \ldots, c_d)$, where $\sum_{i=1}^{d} c_i s_i \leq b$. Of particular interest are those vectors that give rise to a sum of exactly $b$, which we shall call *perfect packing configurations.* For instance, if $S = \{1, 2, 3\}$ and $b = 7$, one such configuration would be $(1, 0, 2)$. Let $\mathbb{P}_{S,b}$ denote the set of all perfect packing configurations for a given $S$ and $b$. Let $\Lambda_{S,b}$ be the convex cone in $\mathbb{R}^d$ spanned by all nonnegative linear combinations of configurations in $\mathbb{P}_{S,b}$.

THEOREM (Courcoubetis and Weber [11]). *For any discrete distribution $F = (b, S, \vec{p})$, the following hold.*

(a) *If $\vec{p}$ lies in the interior of $\Lambda_{S,b}$, then $EW_n^{OPT}(F) = O(1)$.*

(b) *If $\vec{p}$ lies on the boundary of $\Lambda_{S,b}$, then $EW_n^{OPT}(F) = \Theta(n^{1/2})$.*

(c) *If $\vec{p}$ lies outside of $\Lambda_{S,b}$, then $EW_n^{OPT}(F) = \Theta(n)$.*

In general it is NP-hard to determine which of the three cases applies to a given distribution (as can be proved by a straightforward transformation from the PARTITION problem [13]). However, for the distributions $U\{j,k\}$, $j < k-1$, we can use the following lemma, which we shall prove using the perfect packing theorem, to show that (a) applies.

LEMMA 7. *For each $i,j,k$ with $1 \le i \le j < k-1$, there exist positive integers $r_i, s_i, m_i < 2k^2$ such that the set of $r_i j + s_i$ items consisting of $r_i + s_i$ items of size $i$ together with $r_i$ items of each of the other $j-1$ sizes can be packed perfectly into $m_i$ bins of size $k$.*

Note that this lemma implies that the $j$-dimensional vector $\bar{e} = (1/j, 1/j, \ldots, 1/j)$ is strictly inside the appropriate cone when $S = \{1,2,\ldots,j\}$, $j < k-1$. This is because $\bar{e}$ is in the interior of the cone spanned by vectors of the form $(r_i, \ldots, r_i, r_i + s_i, r_i, \ldots, r_i)$, $i = 1, \ldots, j$, and those vectors are sums of perfect packing configurations by Lemma 7. The proof of Theorem 2 thus follows from case (a) of the above theorem.

*Proof of Lemma* 7. We make use of the perfect packing theorem. There are two cases. If $k \ge i+j$, we simply set $r_i = k-i$ and $s_i = m_i = j(j+1)/2$. Note that the total size of $r_i$ items each of the sizes $1,\ldots,j$ equals $r_i j(j+1)/2$, so by the perfect packing theorem, we can perfectly pack them into $j(j+1)/2 = m_i$ bins of size $r_i = k-i$. The remaining $s_i = m_i$ items of size $i$ can then go one per bin to fill these bins up to size precisely $k$.

On the other hand, suppose $k < i+j$. Now things are a bit more complicated. We have $r_i = 2(k-i)$, $s_i = (k-j)(k-j-1)$, and $m_i = s_i + r_i(2j-k+1)/2$. By the perfect packing theorem $r_i$ items each of the sizes $1,\ldots,k-j-1$ perfectly pack in $s_i$ bins of size $k-i$. (Such items exist because by assumption $j < k-1$.) We then add the additional $s_i$ items of size $i$ to these bins, one per bin, to bring each bin up to size $k$. There remain $r_i$ items each of sizes $k-j$ through $j$, for a total of $r_i(j-(k-j-1)) = 2(m_i - s_i)$ items. These can be used to completely fill the remaining $m_i - s_i$ bins with pairs of items of sizes $(j, k-j), (j-1, k-j+1), \ldots, (\lceil k/2 \rceil, \lfloor k/2 \rfloor)$. Note that if $k$ is even, the last bin type contains two items of size $k/2$, but we have an even number of such items by our choice of $r_i = 2(k-i)$, so this presents no difficulty.

It is easy to verify that in both cases $r_i$, $s_i$, and $m_i$ are all less than $2k^2$. □

**4. Proof of Theorem 3.** Recall the theorem statement: If $L_n$ has item sizes generated according to $U(0,u]$ for $0 < u \le 1$, and $A$ is any online algorithm, then there exists a constant $c > 0$ such that $E[W^A(L_n)] > cn^{1/2}$ for infinitely many $n$.

*Proof.* Let $w(t)$ denote the amount of empty space in partially filled bins after $t$ items have been packed. We show that for any $n > 0$ the expected value of the average of $w(1), \ldots, w(n)$ is $\Omega(n^{1/2}u^3)$. This implies that $E[w(n)]$ must be $\underline{\Omega}(n^{1/2}u^3)$, i.e., not $o(n^{1/2})$.

Consider packing item $a_{t+1}$. Let $\upsilon(t)$ denote the number of nonempty bins that have a gap of at least $u^2/8$ after the first $t$ items have been packed. There are at most $\upsilon(t)$ bins into which one can put an item larger than $u^2/8$. Therefore, if $a_{t+1}$ is to leave a gap of less than $\delta$ in its bin, either it must have size less than $u^2/8$ or its size must be within $\delta$ of the empty space in one of these $\upsilon(t)$ bins with gaps larger than $u^2/8$. The probability of this is at most $[u^2/8 + \delta\upsilon(t)]/u$. By choosing $\delta = u^2 n^{-1/2}/8$, conditioning on whether $\upsilon(t)$ is greater or less than $n^{1/2}$, and noting that the size of $a_{t+1}$ is distributed as $U(0,u]$ independent of $\upsilon(t)$, we have

$$P(a_{t+1} \text{ leaves gap } < \delta) \le P(\upsilon(t) \ge n^{1/2}) + u/4.$$

Now

$$\begin{aligned}
E[w(t)] &\ge \delta \sum_{s=0}^{t-1} P(a_{s+1} \text{ is last in a bin and leaves gap } \ge \delta) \\
&= \delta \sum_{s=0}^{t-1} [P(a_{s+1} \text{ is last in a bin}) \\
&\qquad\qquad - P(a_{s+1} \text{ is last in a bin and leaves gap } < \delta)] \\
&\ge \delta \sum_{s=0}^{t-1} [P(a_{s+1} \text{ is last in a bin}) - P(a_{s+1} \text{ leaves gap } < \delta)] \\
&\ge \delta \sum_{s=0}^{t-1} P(a_{s+1} \text{ is last in a bin}) - \sum_{s=0}^{t-1} P(\upsilon(s) \ge n^{1/2}) - \sum_{s=0}^{t-1} u/4.
\end{aligned}$$

Let $S_t$ be the sum of the first $t$ item sizes, and note that $S_t$ is a lower bound on the number of bins and hence on the number of items that are the last item in a bin. We thus have

$$E\left[\sum_{s=0}^{t-1} P(a_{s+1} \text{ is last in a bin})\right] \ge E[S_t] = tu/2.$$

Using the fact that $\delta = u^2 n^{-1/2}/8$, we then have

$$E[w(t)] \ge (u^2 n^{-1/2}/8)\left[tu/4 - \sum_{s=0}^{t-1} P(\upsilon(s) \ge n^{1/2})\right].$$

If $\sum_{s=0}^{n-1} P(\upsilon(s) \ge n^{1/2}) \le nu/24$, we have for all $t \ge n/2$,

$$E[w(t]) \ge (u^2 n^{-1/2}/8)[nu/8 - nu/24] = u^3 n^{1/2}/96.$$

This implies

$$E\left[\frac{1}{n}\sum_{t=1}^{n} w(t)\right] \ge u^3 n^{1/2}/192.$$

On the other hand, if $\sum_{s=0}^{n-1} P(\upsilon(s) \ge n^{1/2}) \ge nu/24$, then

$$\begin{aligned}
E\left[\frac{1}{n}\sum_{t=1}^{n} w(t)\right] &\ge \frac{1}{n}\sum_{t=1}^{n} P(\upsilon(t) \ge n^{1/2}) n^{1/2}(u^2/8) \\
&\ge n^{1/2}(u^2/8)(u/24) = u^3 n^{1/2}/192.
\end{aligned}$$

These imply that $E[w(n)]$ is $\underline{\Omega}(n^{1/2})$. □

It should be noted that the above proof relies heavily on the fact that the distribution is continuous, since this is the reason why the union of $n^{1/2}$ intervals of size $\delta$ cannot cover the full probability space. Our discrete distributions $U\{j,k\}$ do not have this failing, and for this reason we can obtain significantly better average-case behavior for them.

**Table 1** *Expected waste in the symmetric case.*

| | $U(0,1]$ | Ref | $U\{j,k\}$, $j=k-1,k$ | Ref |
|---|---|---|---|---|
| $OPT$, $FFD$, $BFD$ | $\Theta(n^{1/2})$ | [20, 21] | $\Theta(n^{1/2})$ | [10] |
| $SS$ | – | – | $\Theta(n^{1/2}k^{1/2})^{\dagger}$ | [12] |
| $FF$ | $\Theta(n^{2/3})$ | [9] | $\Theta(n^{1/2}k^{1/2})$, $k=O(n^{1/3})$ | [9] |
| | | | $\Theta(n^{2/3})$, $k=\Omega(n^{1/3})$ | [9] |
| $BF$ | $\Theta(n^{1/2}\log^{3/4} n)$ | [26] | $\Theta(n^{1/2}\log^{3/4} k)$ | [6] |
| Best online | $\underline{\Theta}(n^{1/2}\log^{1/2} n)$ | [26, 27] | $\Theta(n^{1/2}\log^{1/2} k)^{\ddagger}$ | |

† The upper bound is proved in the reference; the lower bound is conjectured based on experiments.
‡ The upper and lower bounds here appear to follow from the corresponding results for the continuous case, but the details of the upper bound in particular still need to be worked out.

**5. Concluding Remarks.** The results in this paper were among the first to be obtained about the average-case behavior of bin packing algorithms under discrete distributions. Since they were announced in [4], many additional results have been proved, illustrating further contrasts with (and similarities to) the case of continuous distributions. In this concluding section we survey the literature and point out some of the remaining open problems.

Let us begin by considering symmetric uniform distributions, as represented by $U(0,1]$ and $U\{k-1,k\}$, $k \geq 1$. (In general, a symmetric distribution is one that satisfies $p(s \leq \Delta) = p(s \geq b - \Delta)$ for all $\Delta$, $0 \leq \Delta \leq b$.) Table 1 summarizes what is known about average-case behavior under these distributions. A horizontal line separates the offline algorithms from the online ones. Except where noted, all results in this table are theorems.

Four famous classical algorithms have been extensively studied. *First fit* ($FF$) is an online algorithm in which each item is placed in the first bin that has room for it, where bins are sequenced according to the order in which they received their first item. If no bin has room, a new bin is started. *Best fit* ($BF$) is similar, except now the item is placed in the bin with the smallest gap large enough to contain it (ties broken in favor of the earlier bin). *First fit decreasing* ($FFD$) and *best fit decreasing* ($BFD$) are the corresponding offline algorithms in which the list is first sorted so that the items are in nonincreasing order by size, and then $FF$ ($BF$) is applied. From a worst-case point of view, $FF$ and $BF$ are equivalent: in an asymptotic sense each can produce packings that use 70% more bins than optimal, but neither can do any worse [15]. The corresponding offline versions $FFD$ and $BFD$ each can use $2/9 = 22.22..\%$ more bins than optimal but can do no worse [14, 15].

The results in Table 1 show that these algorithms perform much better on average than in the worst case, since they now have sublinear expected waste, a surprise when it was first observed empirically in [2]. The offline versions continue to have an advantage over their online counterparts, but it is of reduced practical significance. And now there is a distinction in the behavior of $FF$ and $BF$, with $BF$ being the better of the two.

The above remarks apply equally well to the discrete and continuous cases. As to the comparison between these cases, we once again have a significant difference for online algorithms. For any fixed value of $k$, the online algorithms in the table all have $\Theta(n^{1/2})$ expected waste, in contrast to the expected wastes in the continuous case of $\Theta(n^{2/3})$ for $FF$, $\Theta(n^{1/2}\log^{3/4} n)$ for $BF$ and $\underline{\Theta}(n^{1/2}\log^{1/2} n)$ for the best possible online algorithm. (Here the notation $\underline{\Theta}(f(n))$ means that the lower bound is taken in the Hardy and Littlewood sense of "not $o(f(n))$," i.e., "greater than $cf(n)$ for some

**Table 2** *Possibilities for expected waste in the nonsymmetric case.*

| | $U(0,u]$, $u<1$ | Ref | $U\{j,k\}$, $1 \le j \le k-2$ | Ref |
|---|---|---|---|---|
| $OPT$ | $\Theta_u(1)$ | [3] | $\Theta_k(1)$ | [•] |
| $FFD$, $BFD$ | $\Theta_u(1)$, $u \le 1/2$ | [3, 16] | $\Theta_k(1)$, $\Theta_k(n^{1/2})^\dagger$, $\Theta_k(n)$ | [5] |
| | $\Theta_u(n^{1/3})$, $u > 1/2$ | [3, 16] | | |
| $FF$, $BF$ | $\Theta_u(n)^\ddagger$ | [4] | $\Theta_k(1)$, $\Theta_k(n)$, ... ? | [1, 7, 8, 19] |
| Best online | $\underline{\Omega}_u(n^{1/2})$, $O_u(n^{1/2}\log^{3/4} n)$ | [•] [25] | $\Theta_k(1)$ | [•] + [12] |

• Results proved in this paper.
† $\Theta_k(n^{1/2})$ is not ruled out by theorems, but no occurrences are known either. For $BFD$ and $FFD$, it does not occur for any $k \le 10{,}000$ [5].
‡ This is conjectured to hold for all $u \in (0,1)$, based on experimental studies. To date it has been proved only for $u \in [0.66, 2/3)$ and $BF$ [18].

$c > 0$ and infinitely many $n$," rather than in the standard Knuthian sense of "greater than $cf(n)$ for some $c > 0$ and all sufficiently large $n$.")

In a sense, however, the online results for the discrete case are consistent with those for the continuous one. Although technically an online algorithm is not allowed to know the magnitude of $n$, if one formally sets $k = \Theta(n)$ in the formulas for expected waste for the discrete case, one gets $EW_n^A(U\{k-1,k\}) = EW_n^A(U(0,1])$ for $FF$, $BF$, and the best possible online algorithm. $SS$ is not applicable to continuous distributions, but note that in this asymptotic discrete sense it appears to be worse than $FF$ and $BF$. Indeed, experiments suggest that $EW_n^{SS}(U\{n-1,n\}) = \Theta(n)$.

Let us now turn to the nonsymmetric distributions $U(0,u]$, $u < 1$, and $U\{j,k\}$, $j < k-1$. The known results for these distributions are summarized in Table 2. Here for the first time we see differences between the continuous and discrete cases for *offline* algorithms. In particular, for $A \in \{FFD, BFD\}$, $EW_n^A(U(0,u]) = O(1)$ for all $u \le 1/2$ [3, 16], but for many of the distributions $U\{j,k\}$ with $j \le k/2$ (the corresponding discrete uniform distributions), we have $EW_n^A(U\{j,k\}) = \Theta(n)$ [4, 5]. Moreover, for $u \in (1/2, 1)$, $EW_n^A(U(0,u]) = \Theta(n^{1/3})$, and this growth rate *never* occurs for $U\{j,k\}$. This follows from a theorem in [5] that says that for all discrete distributions $F$, $EW_n^{FFD}(F)$ and $EW_n^{BFD}(F)$ must be either $O(1)$, $\Theta(n^{1/2})$, or $\Theta(n)$.

The theorem also provides algorithms that determine the answers for a given distribution $(b, S, \vec{p})$, find the constants of proportionality when the expected waste is linear, and run in time polynomial in $b$ and $|S|$. Unfortunately, although the answers for the distributions $U\{j,k\}$ with $k \le 10{,}000$ have all been computed [5], these do not suggest any simple rule as to how the choice among $O(1)$, $\Theta(n^{1/2})$, and $\Theta(n)$ might depend on $j$ and $k$. (As an example of the type of behavior that can occur, for the $U\{j,151\}$ distributions the choice between linear and bounded expected waste switches back and forth 10 times as $j$ increases from 1 to 149.) The results for $k \le 10{,}000$ do, however, exhibit several suggestive patterns. First (and this can be proved to hold for arbitrarily large $k$), the expected waste is $O(1)$ whenever $j < \sqrt{k}$ or $j > k - \sqrt{k}$. Second, expected waste $\Theta(n^{1/2})$ does not occur for any $U\{j,k\}$ with $j < k-1 < 10{,}000$, suggesting that it may never occur. Third, for each $U\{j,k\}$ with $k \le 10{,}000$, $EW_n^{FFD}(U\{j,k\})$ and $EW_n^{BFD}(U\{j,k\})$ are either both linear or both bounded. If linear, the constant of proportionality for $BFD$ is never larger than that for $FFD$ (but is occasionally smaller).

Here again there is a sense in which the discrete case is asymptotically consistent with the continuous case, even though the expected waste for $FFD$ and $BFD$ is always sublinear in the latter. As $k$ increases, the maximum constants of proportionality for the linear expected waste under $U\{j,k\}$ appear to decrease. Indeed, it can be shown that the constants for $FFD$ are bounded by a function that declines at least as fast as $(\log k)/k$ [5]. (This presumably holds for $BFD$ as well.) The worst case is the distribution $U\{6, 13\}$, for which the expected waste for both $FFD$ and $BFD$ is $n/624$, which is less than 0.6% of the expected optimal number of bins. Moreover, this is easily avoided, since not only does $SS$ have bounded expected waste for this distribution, but so do $FF$ and $BF$ (although this is the only case we have identified where the online $FF$ and $BF$ algorithms outperform their offline cousins).

Turning now to the online algorithms $FF$ and $BF$, we observe that their behavior under discrete uniform distributions appears empirically to be similar to their behavior under continuous ones. Based on extensive experiments, it is conjectured that $FF$ and $BF$ both have linear expected waste under $U(0, u]$ for $0 < u < 1$, although to date this has been proved only for $u \in [0.66, 2/3)$ and $BF$ [18]. In the discrete case ($U\{j, k\}$ with $j < k - 1$) experiments suggest that for sufficiently large $k$, the expected waste for $FF$ and $BF$ is bounded when $j = O(\sqrt{k})$ and when $j = k - 2$, but otherwise linear. Some of this has been proved. In [19] it was shown that $EW_n^{BF}(U\{k - 2, k\}) = O(1)$ for all $k > 0$, and this result was extended to $FF$ in [1]. In [4] it was shown that $EW_n^{FF}(U\{j, k\}) = O(1)$ if $k \geq j(j + 3)/2$ and $EW_n^{BF}(U\{j, k\}) = O(1)$ if $k \geq j^2$. In practice, bounded expected waste is more common, at least for small $k$. The growth rates for $BF$ under $U\{j, k\}$ with $k \leq 11$ were completely characterized using multidimensional Markov chain arguments in [8], and linear expected waste only occurs for $U\{8, 11\}$. The only general result proving linear expected waste mirrors the result for the continuous case: $EW_n^{BF}(U\{j, k\}) = \Theta(n)$ if $j/k \in [.66, 2/3)$ and $k$ is sufficiently large [18]. At this point we do not know if expected wastes other than $O(1)$ and $\Theta(n)$ are possible for $FF$ or $BF$ under any distributions $U\{j, k\}$ with $j < k - 1$. No general classification theorem such as those for $FFD$, $BFD$, or $OPT$ has been proven, so the range of possibilities is not known to be limited to $O(1)$, $\Theta(n^{1/2})$, and $\Theta(n)$, as it was for $FFD$ and $BFD$.

There is also a gap between the lower bound proved in this paper on the best possible online expected waste for continuous distributions $U(0, u]$ and the best rate known to be achievable. The former is $\underline{\Omega}(n^{1/2})$ and the latter is $O(n^{1/2} \log^{3/4} n)$, as proved in [25]. The algorithm of [25] works for any distribution, discrete or continuous, but has drawbacks from a pragmatic point of view: the best current bound we have on its running time is $O(n^8 \log^3 n)$ [12]. If one is willing to consider more specialized algorithms, better running times are possible, at least theoretically. For any fixed distribution $F$, there is an algorithm $A_F$ that runs in time $O(n \log n)$ and again has expected waste of $O(n^{1/2} \log^{3/4} n)$ [24]. These algorithms have drawbacks too, however, since the proof that they exist is nonconstructive. The question of whether practical algorithms exist that attain these bounds, or indeed whether the $\underline{\Omega}(n^{1/2})$ lower bound is achievable, remains open.

Finally, in addition to the open problems mentioned above for the discrete and continuous uniform distributions $U(0, u]$ and $U\{j, k\}$, there is the question of what happens for arbitrary discrete and continuous distributions. In the discrete case, the above-mentioned classification theorems apply for $BFD$, $FFD$, and $OPT$, and say that the corresponding expected waste must be $O(1)$, $\Theta(n^{1/2})$, or $\Theta(n)$. As also mentioned above, the applicable cases for $BFD$ and $FFD$ and any specific distribution $F = (b, S, \vec{p})$ can be determined in time polynomial in $b$ and $|S|$. For $OPT$ there is also an algorithm for determining which case applies, as noted in [12]. This involves solving up to $|S| + 1$ linear programs with $|S|b$ variables and $|S| + b$ constraints. None of these algorithms technically runs in polynomial time since $b$ may be exponentially larger than its contribution to instance size ($\log b$). However, all are feasible for $b$ in excess of 1,000, which makes it possible to characterize the behavior of $FFD$, $BFD$, and $OPT$ for many interesting distributions on a case-by-case basis.

The theorem about $SS$ presented in section 1 can be generalized to arbitrary discrete distributions if one replaces $SS$ by a simple variant $SS'$: As in $SS$, items are packed so as to minimize $\sum_{h=1}^{b-1} N_h^2$, but now the choice must be made subject to the following additional constraint: No item may be placed in a partially filled bin if the resulting gap cannot be exactly filled with items whose sizes have already been encountered in the list $L$. The resulting algorithm still runs in time $O(nb)$ and satisfies $EW_n^{SS'}(F) = \Theta\left(EW_n^{OPT}(F)\right)$ for all discrete distributions $F$ [12]. In addition, there is a more complicated randomized variant that runs in time $O(nb \log b)$, satisfies the above property, and also has the same constant of proportionality as $OPT$ when the expected waste is linear [12].

As to the case of arbitrary continuous distributions, we as yet have no general classification theorems, although some partial results have been proved. Rhee [22] provided a complicated measure-theoretic characterization of those $F$ for which $EW_n^{OPT}(F)$ is sublinear, but this does not appear to be computationally useful. A result of Rhee and Talagrand [23] implies that if $EW_n^{OPT}(F)$ is sublinear, it must be $O(n^{1/2})$ or better. Rates strictly between $O(1)$ and $\Theta(n^{1/2})$ have not yet been ruled out, however. Moreover, there is as yet no algorithm with the general effectiveness of $SS$ and its variants. The results of [24, 25] imply that there are online algorithms whose expected waste is at most $O(n^{1/2} \log^{3/4} n)$ worse than the optimal expected waste. For offline algorithms, Karmarkar and Karp have devised an algorithm which in the *worst case* never uses more than the optimal number of bins plus $O\left(\log^2(OPT)\right) = O(\log^2 n)$ [17]. This means that its expected waste is never more than the maximum of $O(\log^2 n)$ and $EW_n^{OPT}(F)$. Like the algorithms of [24, 25], however, it is impractical, having a running time for which our best current bound is $O(n^8 \log^2 n)$.

To conclude with an open problem that hearkens back to the main result of this paper, note that our ability to determine the expected waste for $FFD$, $BFD$, and $OPT$ on a case-by-case basis can only take us so far, and more general results would be desirable. Results for $U(0,1]$ and $U\{k-1,k\}$ typically continue to hold for arbitrary continuous and discrete symmetric distributions, respectively, but the real world is not dominated by symmetric distributions. It would be nice if we could identify additional interesting classes of nonsymmetric distributions $F$ for which general results about $EW_n^{OPT}(F)$ can be proved, as we did in this paper for the discrete uniform distributions. Are there interesting classes for which new perfect packing theorems can provide us with similar general answers?

## REFERENCES

[1] S. Albers and M. Mitzenmacher, *Average-case analysis of first fit and random fit bin packing*, in Proceedings of 9th ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 1998, pp. 290–299.

[2] J. L. Bentley, D. S. Johnson, F. T. Leighton, and C. C. McGeoch, *An experimental study of bin packing*, in Proceedings of the 21st Annual Allerton Conference on Communication, Control, and Computing, University of Illinois, Urbana, 1983, pp. 51–60.

[3] J. L. Bentley, D. S. Johnson, F. T. Leighton, C. C. McGeoch, and L. A. McGeoch, *Some unexpected expected behavior results for bin packing*, in Proceedings of the 16th Annual ACM Symposium on Theory of Computing, ACM Press, New York, 1984, pp. 279–288.

[4] E. G. Coffman, Jr., C. Courcoubetis, M. R. Garey, D. S. Johnson, L. A. McGeoch, P. W. Shor, R. R. Weber, and M. Yannakakis, *Fundamental discrepancies between average-case analyses under discrete and continuous distributions*, in Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, ACM Press, New York, 1991, pp. 230–240.

[5] E. G. Coffman, Jr., D. S. Johnson, L. A. McGeoch, P. W. Shor, and R. R. Weber, *Bin Packing with Discrete Item Sizes, Part* III: *Average Case Behavior of FFD and BFD*, manuscript.

[6] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber, *Bin Packing with Discrete Item Sizes, Part* V: *Tight Bounds on Best Fit*, manuscript.

[7] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber, *Bin Packing with Discrete Item Sizes, Part* IV: *Markov Chains, Computer Proofs, and Average-Case Analysis of Best Fit Bin Packing*, manuscript.

[8] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber, *Markov chains, computer proofs, and average-case analysis of Best Fit bin packing*, in Proceedings of the 25th Annual ACM Symposium on Theory of Computing, ACM Press, New York, 1993, pp. 412–421.

[9] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber, *Bin packing with discrete item sizes, Part* II: *Tight bounds on First Fit*, Random Structures Algorithms, 10 (1997), pp. 69–101.

[10] E. G. Coffman, Jr. and G. S. Lueker, *An Introduction to the Probabilistic Analysis of Packing and Partitioning Algorithms*, Wiley, New York, 1991.

[11] C. Courcoubetis and R. R. Weber, *Stability of on-line bin packing with random arrivals and long-run average constraints*, Probab. Engrg. Inform. Sci., 4 (1990), pp. 447–460.

[12] J. Csirik, D. S. Johnson, C. Kenyon, J. B. Orlin, P. Shor, and R. R. Weber, *On the sum-of-squares algorithm for bin packing*, in Proceedings of the 32nd ACM Symposium on Theory of Computing, ACM Press, New York, 2000, pp. 208–217.

[13] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, W. H. Freeman, New York, 1979.

[14] D. S. Johnson, *Near-Optimal Bin Packing Algorithms*, Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 1973.

[15] D. S. Johnson, A. Demers, J. D. Ullman, M. R. Garey, and R. L. Graham, *Worst-case performance bounds for simple one-dimensional packing algorithms*, SIAM J. Comput., 3 (1974), pp. 299–325.

[16] D. S. Johnson, F. T. Leighton, R. R. Weber, and P. W. Shor, *Average Case Behavior of First Fit Decreasing and Optimal Packings for Continuous Uniform Distributions* $U(0,u]$, manuscript.

[17] N. Karmarkar and R. M. Karp, *An efficient approximation scheme for the one-dimensional bin packing problem*, in Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1982, pp. 312–320.

[18] C. Kenyon and M. Mitzenmacher, *Linear waste of best fit bin packing on skewed distributions*, in Proceedings of the 41st Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 2000, pp. 582–589.

[19] C. Kenyon, Y. Rabani, and A. Sinclair, *Biased random walks, Lyapunov functions, and stochastic analysis of best fit bin packing*, J. Algorithms, 27 (1998), pp. 218–235.

[20] W. Knödel, *A bin packing algorithm with complexity* $o(nlogn)$ *in the stochastic limit*, in Proceedings of the 10th Symposium on Mathematical Foundations of Computer Science, J. Gruska and M. Chytil, eds., Lecture Notes in Comput. Sci. 118, Springer-Verlag, Berlin, 1981, pp. 369–378.

[21] G. S. Lueker, *An Average-Case Analysis of Bin Packing with Uniformly Distributed Item Sizes*, Tech. Report 181, Department of Information and Computer Science, University of California, Irvine, CA, 1982.

[22] W. T. Rhee, *Optimal bin packing with items of random sizes*, Math. Oper. Res., 13 (1988), pp. 140–151.

[23] W. T. Rhee and M. Talagrand, *Optimal bin packing with items of random sizes.* III, SIAM J. Comput., 18 (1989), pp. 473–486.

[24] W. T. Rhee and M. Talagrand, *On-line bin packing with items of random size*, Math. Oper. Res., 18 (1993), pp. 438–445.

[25] W. T. Rhee and M. Talagrand, *On-line bin packing of items of random sizes.* II, SIAM J. Comput., 22 (1993), pp. 1251–1256.

[26] P. W. Shor, *The average case analysis of some on-line algorithms for bin packing*, Combinatorica, 6 (1986), pp. 179–200.

[27] P. W. Shor, *How to pack better than Best Fit: Tight bounds for average-case on-line bin packing*, in Proceedings of the 32nd Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1991, pp. 752–759.